Fluorescent amino acids as versatile building blocks for chemical biology

Fluorescent amino acids as versatile building blocks for chemical biology

Zhiming Cheng[1], Erkin Kuru[2], Amit Sachdeva[3], Marc Vendrell[1†]

[1]Centre for Inflammation Research, University of Edinburgh, Edinburgh, UK
[2]Department of Genetics, Harvard Medical School, Boston, USA
[3]School of Chemistry, University of East Anglia, Norwich, UK

[†]email: marc.vendrell@ed.ac.uk

Abstract | Fluorophores have transformed the way we study biological systems, enabling non-invasive studies in cells and intact organisms, which increase our understanding of complex processes at the molecular level. Fluorescent amino acids (FlAAs) have become an essential chemical tool because they can be used to construct fluorescent macromolecules, such as peptides and proteins, without disrupting their native biomolecular properties. Fluorescent and fluorogenic amino acids with unique photophysical properties have been designed for tracking protein–protein interactions *in situ* or imaging nanoscopic events in real-time with high spatial resolution. In this Review, we discuss advances in the design and synthesis of FlAAs and how they have contributed to the field of chemical biology in the past 10 years. Important areas of research that we review include novel methodologies to synthesize building blocks with tunable spectral properties, their integration into peptide and protein scaffolds using site-specific genetic encoding and bio-orthogonal approaches, and their application to design novel artificial proteins as well as to investigate biological processes in cells by means of optical imaging.

[H1] Introduction

Fluorescence-based techniques have revolutionized our understanding of biological processes, as they allow researchers to examine the localization, trafficking and activity of biomolecules in cells, tissues and even whole intact organisms in a non-invasive manner.[1-4] Many natural peptides and proteins have key roles in various biological processes; however, directly visualizing these molecules is hampered by the fact they generally do not exhibit strong fluorescence emission. Whereas this low 'background' fluorescence boosts the sensitivity of spectroscopy and imaging experiments, its weak levels highlight the need for labelling

strategies that can facilitate the detection of these macromolecules. Fluorescent labelling of peptides and proteins has been successfully used in numerous ways, from the widely used expression of fluorescent protein fusions (such as green fluorescent protein[5] (GFP) and subsequent derivatives[6,7]) to the fusion of proteins with self-labelling tags (such as SNAP, CLIP and HALO tags[8-10]) or modification with specific peptide sequences that can bind to small molecules with high affinity (such as FlASH and ReASH[11,12]). All of these approaches have led to remarkable biological discoveries, but they typically involve a large structural modification of the peptides or proteins of interest, which can alter their biomolecular properties.

Non-natural fluorescent amino acids (FlAAs) have been developed as chemical alternatives for the derivatization of peptides and proteins in a less disruptive manner. Owing to their small size and similarity to the natural residues found in polypeptides, they can be used to fluorescently label macromolecules while retaining their overall function with minimal perturbation of the native protein structure. The toolbox of FlAAs expanded considerably in the early 2000s, partly due to the development of numerous FlAAs[13] with unique optical properties, including environmental sensitivity[14-16], responsive to metal chelation[17], tunable fluorescence emission and prolonged lifetime[18]). Importantly, these modular building blocks can be introduced at specific sites in small peptides using solid-phase peptide synthesis[19] (SPPS) or into larger proteins by genetic encoding[20] to generate fluorescent, native-like biomolecules, enabling biological experiments that would be intractable by other means, including analysis of protein conformational changes, monitoring of protein–protein interactions and activity studies in live cells.

Over the past 10 years, a substantial increase in the number of versatile synthetic strategies (such as multicomponent reactions[21], metal-catalysed reactions[22], light-induced transformations[23] and bioorthogonal chemistry[24]) to fine-tune the chemical structure of fluorophores[25] and natural building blocks has accelerated the design and preparation of new FlAAs. In this Review, we provide an overview of different chemical transformations that have been used to generate novel building blocks with bespoke fluorescent properties and to facilitate their site-specific integration into peptides and proteins. We also discuss how the range of applications of FlAAs has extended to include many biological assays *in vitro*, in cells and *in vivo*, which has been facilitated by a rapid expansion in the use of spectroscopic and imaging-based techniques. Finally, we review the versatility of FlAAs from the point of view of an end-user and present several examples of the application of FlAAs in chemical biology,

including molecular recognition studies, live-cell imaging and applications in synthetic biology and microbiology (**Figure 1**).

[H1] Synthesis of FlAAs

[H2] Fluorescent analogues of natural amino acids

Three canonical amino acids, (tryptophan, tyrosine and phenylalanine, are fluorescent and thus might have potential as naturally-occurring fluorophores[26]; however, their optical properties (such as excitation and emission wavelengths, brightness and photostability) are suboptimal for most biological assays. Tryptophan, the most commonly used fluorophore among natural amino acids, absorbs and emits in the ultraviolet range (absorption wavelength (λ_{abs}) = 280 nm; emission wavelength (λ_{em}) = 346 nm) and its fluorescence quantum yield is ~20%[27].

Initial attempts to improve the fluorescent properties of tryptophan yielded azatryptophans as potential isosteric substitutes in proteins[28]. Although 4-azatryptophans have a longer emission wavelength ($\lambda_{em.}$= 425 nm) than tryptophan and a large Stokes shift (>130 nm), they are typically less bright than tryptophan. Alternative approaches have involved the preparation of cyanotryptophans, which have improved fluorescence quantum yields (approaching ~50%[29]) owing to the incorporation of the strong electron withdrawing cyano group. Cyanotryptophans can be synthesized from commercially available cyanoindoles, and the red-shifted emission wavelength of 6-cyanotryptophan (**1a**) (λ_{em}= 370 nm) enabled its use for Förster resonance energy transfer (FRET) experiments to study protein–DNA interactions. Subsequent studies identified 4-cyanotryptophan (**1b**) as an amino acid with improved optical properties[30]. A key step in the synthesis of 4-cyanotryptophan is the palladium (Pd)-catalysed incorporation of the cyano group at position 4 of L-tryptophan. Remarkably, the emission maximum of 4-cyanotryptophan is in the blue visible region (λ_{em} = 405 nm) and it has a high quantum yield (~80%), good photostability and a long fluorescence lifetime (~13.7 ns), which makes it an interesting building block for spectroscopic and microscopic measurements of proteins, such as the study of peptide–membrane interactions[31]. The structural similarity of cyanotryptophans to tryptophan also makes them an attractive platform to explore enzyme-based reactions, facilitating the preparation of enantiopure tryptophan analogues by derivatization at multiple aromatic positions, using, for example, tryptophan synthase S (TrpS)[32] or variants of its β-subunit (TrpB) [33].

Extended tryptophan analogues have been also designed by conjugation of homocyclic and heterocyclic systems to the indole moiety, and include tricyclic tryptophan analogues (**2**)

derived from pyrrolo(iso)quinolines[34], tryptophan derivatives obtained by hydrogen-mediated coupling of 1,2,3-triazoles to indoles[35] (**3a**) and a series of 2-arylated tryptophans (**3b**) obtained by Pd-mediated C–H functionalization[36]. Cross-coupling reactions (such as Suzuki–Miyaura couplings) have also been effective for preparing C4-substituted tryptophans[37], although the effect of these substitutions on the optical properties of tryptophan has not been reported. Of note, extended tryptophans have a red-shifted emission wavelength ($\lambda_{em.}$= 420–480 nm)[34-36] and some have been incorporated in *Escherichia coli* dihydrofolate reductase (DHFR) using *in vitro* translation, resulting in minimal disruption of DHFR function[38] and thus demonstrating the potential of these molecules for monitoring conformational changes in macromolecular structures.

[H2] Non-natural FlAAs

Despite the remarkable progress in the synthesis of fluorescent analogues of natural amino acids, numerous non-natural FlAAs with different optical properties have been developed (**Figure 2**). The synthetic approaches to generate non-natural FlAAs can be broadly categorized into those that involve appending fluorescent moieties to amino acids and those that involve *de novo* construction of amino acids with integrated chromophores (**Figure 2**).

[H3] Palladium catalysis. Pd-catalysed reactions are one of the most commonly used approaches for the synthesis of non-natural FlAAs. Biphenyl derivatives of phenylalanine (**4a**) have been prepared by Suzuki coupling of biphenyl boronic acids with iodophenylalanine[39,40]. Their emission maximum is red-shifted relative to that of phenylalanine (~340 nm versus 282 nm, respectively) and they have high quantum yields (~70%). Suzuki couplings have also been used in the synthesis of fluorescent D-amino acids (FDAAs), including 4-acetamidobiphenyl groups (**4b**), which have a slightly longer emission wavelength (~380 nm)[41]. Heck reactions have been employed to extend the π-conjugation of tyrosine[42], yielding styryl-containing tyrosine analogues (**5**) with tunable fluorescence quantum yields (from 10% to 90%) and emission maximathat cover the entire visible spectrum ($\lambda_{em.}$= 400–800 nm). Sonogashira coupling reactions have been used to append aryl-containing acetylene derivatives to benzoxazol-5-yl-alanine to produce FlAAs (**6**) with a high quantum yield and emission maxima in ultraviolet and blue wavelengths ($\lambda_{em.}$= 340–450 nm) [43]. Similarly, xanthone structures ($\lambda_{em.}$= 380 nm) have been incorporated into the alanine core by Negishi couplings[44] (**7a**). The high efficiency of Pd-catalysed couplings (such as the Buchwald–Hartwig reaction) has also

facilitated the synthetic optimization of widely used non-natural FlAAs, such as acridon-2-ylalanine (ACD) (**7b**), which is one of the brightest blue-emitting FlAAs reported to date[45].

[H3] C–H activation. C–H activation is another synthetic methodology that has contributed substantially to the development of non-natural FlAAs, particularly for the construction of new tryptophan-based amino acids. Trp-BODIPY (**8a**) was the first BODIPY-containing tryptophan analogue, which has excellent photophysical properties ($\lambda_{abs.}$= 500 nm; $\lambda_{em.}$= 530 nm) and is compatible with SPPS[46]. A similar approach was used to prepare Trp(redBODIPY) (**8b**), which has a red-shifted maximum excitation and emission wavelength ($\lambda_{abs.}$= 560 nm, $\lambda_{em.}$= 590 nm)[47], and also was extended for the synthesis of styryl-derivatized tryptophan analogues by C–H olefination at the C2 position of the indole moiety[48]. Given that the extension of π-conjugation systems often leads to improved photophysical properties, as described above with Heck reactions[42], this strategy has the potential to yield FlAAs with novel optical properties. Pd-catalysed sp^3 C–H activation has also been used to append green and red-fluorescent BODIPY dyes to the side chains of alanine and phenylalanine (**9**) ($\lambda_{em.}$= 510–625 nm)[49] and to the C2 position of tryptophan by a novel ruthenium(II)-catalysed C–H alkylation[50].

[H3] Copper and rhodium catalysis. Further examples of metal-catalysed reactions include the Ullmann-type reaction for coupling the fluorophore pyrene to the phenol group of tyrosine (**10a**) using copper catalysis ($\lambda_{em.}$= 390 nm; quantum yield ~40%)[51], or the widely used copper-catalysed azide–alkyne cycloaddition (CuAAC) to conjugate standard fluorophores, such as coumarins -($\lambda_{em.}$= 380 nm), benzothiadiazoles ($\lambda_{em.}$= 470 nm), fluorescein ($\lambda_{em.}$= 510 nm), dansyl ($\lambda_{em.}$= 520 nm), nitrobenzodioxazoles (NBDs; $\lambda_{em.}$= 530 nm), and naphthalimides ($\lambda_{em.}$= 550 nm), to the side chains of different amino acids[52-54] (**11**). Fluorescent arylalanines (**12**) ($\lambda_{em.}$= 400 nm; quantum yield ~60%) have been synthesized via the formation of C–C bonds between several polyaromatic hydrocarbons and alanine using rhodium catalysis[55].

[H3] Subheading. Conventional reactions involving the formation of amides[56-61] (**13**), carbamates[62] (**14**) or thioureas[57,58,63] (**15**) have also been explored for modifying the side chains of different amino acids with various fluorophores. Other reactions include Knoevenagel-like condensation to derivatize aminodicarboxylate β-ketoesters with pyrene[64] (**16**) or synthesis of a green-fluorescent FlAA (**17**) ($\lambda_{em.}$= 540 nm) by coupling 4-dimethylamino-1-naphthalenes to methyl (2*S*)-5-(dimethoxyphosphoryl)-4-oxo-2-(tritylamino)-pentanoate via a Horner–Wadsworth–Emmons reaction[65]. The simplicity of some reactions and the availability of properly derivatized fluorescent building blocks has also facilitated the preparation of

collections of FlAAs. For example, using Michael addition, a small library of D-cysteine FlAAs (**18**) with thiol-containing fluorophores was generated[66], and maleimide-4-aminophthalimide[67] (**19**) was coupled to L-cysteine. Furthermore, the Fukuyama–Mitsunobu reaction was used to couple 6-acyl-2-naphthylamine to L-serine to produce 3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (ANAP) (**20**), a building block with remarkable optical properties ($\lambda_{em.}$= 490 nm, quantum yield ~50%) that is widely used for fluorescent labelling of proteins[68]. NBD-like fluorophores have been conjugated by a nucleophilic aromatic substitution (S_NAr) reaction to nucleophilic groups (**21**) in amino acid side chains (such as 3-amino-alanine)[58,69], whereas phospholyl(borane) amino acids (**22**) have been produced by nucleophilic substitution of fluorescent phospholide anions with iodo-derivatized amino acids[70], which enabled the synthesis of a range of phospholyl derivatives with fluorescence properties covering a broad spectrum of emission maxima wavelengths from ultraviolet to green ($\lambda_{em.}$= 340–530 nm).

[H3] Flavone-based FlAAs. Alternative synthetic approaches have involved the *de novo* design of FlAAs by building fluorophores into an amino acid structure (**Figure 3**). For example, flavone-based amino acids (**23**) (**Figure 3a**) are produced by constructing the fluorophore 3-hydroxychromone from tyrosine by aldol condensation, followed by oxidative cyclization (for example, the Algar–Flynn–Oyamada reaction)[71-73]. Of note, flavone FlAAs show excited-state intramolecular proton transfer (ESIPT) fluorescence, and they only minimally perturb the structure of peptides owing to their small size. Another FlAA with a built-in fluorophore, benzoacridone-modified alanine, was synthesized by an Ullmann-type coupling between 3-chloro-2-naphthoic acid and aminophenylalanine, followed by cyclization of 2-aminonapthoic acid to yield green-fluorescent benzoacridone amino acid ($\lambda_{em.}$= 500–550 nm, quantum yield ~50%)[74]. Furthermore, a non-natural FlAAs have been synthesised by building aryl-substituted pyridyl, pyrazole, benzotriazole and pyrazoloquinazoline (**24**) heterocyclic chromophores on aspartic acid and asparagine structures ($\lambda_{em.}$= 348–460 nm)[75-78] (**Figure 3b**).

[H3] Coumarin based FlAAs. The simplicity of the coumarin scaffold (that is, 1-benzopyran-2-one) has facilitated the preparation of FlAAs integrating this fluorescent structure. Coumarin-based FlAAs can be prepared by Pechmann condensation, either between an amino acid containing a β-ketoester unit and a phenol derivative or between tyrosine and an ethyl acetoacetate derivative[79,80]. Alternatively, 7-(hydroxycoumarin-4-yl)ethylglycine is prepared by Pd-catalysed cross coupling between 2,4-dimethoxyphenylboronic acid and amino acids containing β-ketoester groups[81]. The brightness of coumarin-based FlAAs, with quantum

yields reaching 70%, together with their fairly small size, have made them excellent building blocks for multiple applications in chemical biology.

[H1] Bioactive fluorescent peptides

[H2] Studying biomolecular interactions

Peptides are excellent scaffolds for biological studies, as they can be used to monitor highly specific molecular interactions with a broad range of biomolecules. Fluorescent peptides have traditionally been synthesized by coupling fluorophores to reactive groups (initially to amines, carboxylic acids and thiols[82], but now also to imidazoles[83] or phenols[84]) on the side chains of amino acids or in conveniently placed spacers. However, in some cases, this synthetic approach can alter the native conformation of the peptide or its biological properties (such as functional activity, binding affinity and/or subcellular localization)[85]. One synthetic alternative to minimize the effect of fluorescent labelling is to embed FlAAs within the sequence of the peptides (to avoid the modification of polar groups, such as amines, carboxylic acids, and thiols, which may be crucial for its bioactivity) or to optimize suitable spacers[86]. These methodologies rely on the chemical robustness and flexibility of SPPS, which enables the efficient preparation of highly diverse peptides, including those that incorporate non-natural FlAAs[87].

[H3] Monitoring protein–protein and protein–DNA interactions. Most biological processes rely on the interaction of proteins with their binding partners. In this context, fluorescent peptides have been developed to better understand protein-based interactions at the molecular level and to help in the discovery of new drugs to modulate them. For example, flavone-based solvatochromic FlAAs were incorporated at specific sites in a peptide sequence from the human immunodeficiency virus-1 (HIV-1) nucleocapsid protein during SPPS[71,73]. These FlAAs have hydration-sensitive dual emission maxima ($\lambda_{em.}$= 430 nm and 530 nm), which allowed the interaction between the nucleocapsid protein and nucleic acids to be monitored to gain insights into the HIV-1 viral life cycle (**Figure 4a**). Other fluorescent structures have been explored as solvatochromic amino acids with high sensitivity to their surrounding microenvironments. For example, phthalimide-based amino acids have been integrated into peptide structures to study dynamic protein–protein interactions, as in the case of the PDZ domain[88] (**Figure 4b**), one of the most abundant protein interaction domains in eukaryotes, or calmodulin[89], a major calcium-signal transduction protein. Subsequently, novel

fluorogenic peptides have been used as conformational reporters of calmodulin activity[90,91]. In this case, calmodulin-binding peptides were first identified by *in vitro* selection using tRNA carrying an NBD-based amino acid[90], which was later replaced with a 4-*N*,*N*-dimethylamino-1,8-naphthalimide amino acid that has a similar emission wavelength ($\lambda_{em.}$= 530 nm) but a remarkably enhanced fluorogenic behaviour (up to 100-fold fluorescence increase after binding to calmodulin) [91]. Furthermore, a fluorogenic peptide containing the environmentally-sensitive 1,5-naphthyridin-2(1*H*)-one amino acid has been used to obtain ratiometric fluorescence measurements at two emission wavelengths (370 nm and 480 nm) in response to conformational changes in calmodulin[61].

[H3] Monitoring peptide-membrane interactions. Fluorogenic peptides can also be designed to study peptide–membrane interactions, which are crucial for peptide-based therapeutics and transfection reagents. Cationic peptides, including antimicrobial peptides, are suitable for this purpose, given their high membrane-binding activity. Fluorogenic analogues of melittin, an active component of honey bee venom, have been used to study the orientation of melittin in lipid-rich membranes, revealing details about the varying insertion depths of different residues (that is, flavone FlAAs) and that melittin is oriented parallel to the surface of cell membranes[72]. Similar approaches have been used to study the membrane binding of neuropeptides, which function as messengers between neurons. Analysis of the fluorescence of two analogues of the insect neuropeptide helicokinin I (containing the FlAAs *β*-(4′-hydroxy-′-benzoyl) (ALB; $\lambda_{em.}$= 360 nm) or 6-dimethyl-amino-2-acyl-naphthalene (ALD; $\lambda_{em.}$= 530 nm)) interacting with dodecylphosphatidylcholine micelles was used to develop a model of helicokinin I binding to membranes[92]. Furthermore, Fmoc-L-4-cyanotryptophan was incorporated in the membrane-interacting pH-(low) insertion peptide (pHLIP) during SPPS to measure peptide–membrane binding constants using FRET experiments with the universal membrane stain DiO (ref.[31]) (**Figure 4c**).

[H3] Analytical applications. Fluorogenic peptides have also been used as analytical tools for the fluorometric detection of specific biomolecules. For example, peptide aptamers containing an environmentally-sensitive NBD FlAA were evolved *in vitro* using ribosome display to identify those that detect verotoxin, a protein that is produced by *E. coli* and is associated with haemolytic uraemic syndrome[93]. Furthermore, a 7-azatryptophan-containing peptide has been used for detection of haeme levels in cells[94] using FRET-based measurements based on the overlap between the emission spectrum of 7-azatryptophan ($\lambda_{em.}$= 400 nm) and the absorbance of the haeme group. Peptides containing other fluorophores have also been used in FRET

assays to measure protease activity[69], including benzoacridone amino acids to detect active caspase 3 (ref.[74]), a key enzyme in the apoptotic pathway.

[H2] Live-cell optical imaging

Optical microscopy has allowed researchers to examine the movement of biomolecules inside the cell with high spatial and temporal resolution. Fluorescent peptides are perfectly suited for optical imaging, as they can target specific proteins in cells and also contain optical reporters (that is, FlAAs) that are easily detected using fluorescence microscopes[95]. To speed up the detection of pulmonary infections, fluorogenic Trp–BODIPY-containing antimicrobial peptides were designed for visualization of the deadly fungal pathogen *Aspergillus fumigatus* in *ex vivo* human lung tissue[96] (**Figure 4d**). This residue is an optimal reporter as it retains the molecular recognition of the native tryptophan and also provides a fluorogenic readout on specific binding to fungal cells. Other fluorogenic antimicrobial peptides have enabled *in situ* detection of bacteria in explanted whole human lungs using real-time optical endomicroscopy[97]. The versatility of Trp-BODIPY has led to the preparation of peptide imaging agents for additional targets. For example, a fluorogenic cyclic peptide that mimics milk fat globule EGF factor 8 (MFGE8; also known as lactadherin), a protein with high binding affinity for phosphatidylserine, has been used to image the release of apoptotic bodies into the extracellular space during programmed cell death[98]. A red analogue, Trp–redBODIPY, has been used to prepare fluorogenic cyclic peptides that bind to keratin 1 (KRT1). These peptides were synthesized using Pd-catalysed C–H activation during SPPS, a labelling approach that is less disruptive than conventional lysine derivatization and enabled imaging studies to understand the interaction between immune cells and cancer cells in aggressive breast cancer tumours[47].

Imaging studies can be also performed to acquire functional readouts from cells. For example, a fluorogenic coumarin-based amino acid ($\lambda_{em.}$= 460 nm), the first phosphotyrosine-mimetic FlAA, was used to report the endogenous phosphatase activity of protein tyrosine phosphatases in live cells[99] (**Figure 4e**). The wash-free imaging capabilities of fluorogenic peptides makes them valuable tools for applications where samples must be rapidly analysed with few processing steps, such as in clinical diagnostics or metabolic engineering. For example, NBD-aminophenylalanine (NBD-amPhe) has been used as a building block for the preparation of epithelial cell adhesion molecule (EpCAM)-binding peptides that can detect

circulating tumour cells in the blood[100]. Furthermore, fluorogenic peptide aptamers have been used to image in real time the production of paramylon (a carbohydrate granule similar to starch) by the microalga *Euglena gracilis*, thereby opening new opportunities in metabolic engineering[101].

[H1] Fluorescent D-amino acids in bacteria

[H3] Visualizing bacterial cell wall growth

Bacterial cells contain two main types of macromolecules that are assembled from amino acids — namely, proteins, which consist of L-amino acids, and peptidoglycans, which contain both L-amino acids and D-amino acids. Peptidoglycans are complex polymers that form the cell wall of bacteria and coordinate multiple important processes, including cell growth and division[102]. Their biological importance has made peptidoglycans the target of many antibiotics, and peptidoglycan biosynthesis has become an area of extensive research for the discovery of new antimicrobial drugs[103]. The discovery of fluorescent D-amino acids (FDAAs) has provided researchers in this field with non-invasive probes to visualize key steps during the peptidoglycan biosynthesis in bacterial cells[104,105].

Taking advantage of the inherent promiscuity of taxonomically-diverse bacterial to incorporate D-amino acids as peptidoglycan metabolites[106-108], modified D-amino acids have been used to specifically label sites of new cell wall growth in real time[109,110]. Structurally-diverse FDAAs, including different reporters of varying size and optical properties, have been synthesised[111] (**Figure 5a**). For example, ethynyl-D-alanine (EDA), azido-D-alanine (ADA) and dipeptides, such as ethynyl-D-alanyl-D-alanine (EDA-DA), have been used to label peptidoglycans in different species of bacteria by cycloadditions with fluorophores. These bioorthogonal approaches require two reaction steps that can compromise cell viability; however, these 'clickable' D-amino acidsare small and thus are minimally disruptive and generally compatible with multiple reporters. For example, the small size and biological stability of AlexaFluor 488-conjugated EDA-DA enabled the demonstration that peptidoglycans are present in the cell wall of the human pathogen *Chlamydia trachomatis*[111]. Alternatively, FDAAs have been employed for single-step labelling of peptidoglycans during their biosynthesis in bacterial cells, for example, by the metabolic incorporation of FDAAs into new peptidoglycans to label bacteria with minimal perturbation of the cells and simpler protocols (that is, fewer washing steps). The good selectivity for bacteria and straightforward application of FDAAs has opened multiple research avenues in different biological

applications, ranging from *in vivo* analysis of the gut microbiota in mice[112] and in human fecal samples[113] to light-induced strategies for killing pathogenic bacteria[114].

[H3] Multi-colour and multiplexed imaging

The simple, modular design of FDAAs and their convenient synthesis by condensation of standard fluorophores to D-amino acids (such as D-lysine and D-diaminopropionic acid), has not only facilitated their diversification with a broad range of fluorophores but has also enabled imaging experiments to answer questions about bacterial growth and division with high spatial resolution. For example, live-cell imaging experiments in which three differently-coloured FDAAs, NBD-3-amino-D-alanine (NADA; $\lambda_{em} = 538$ nm), tetramethylrhodamine-D-lysine (TDL; $\lambda_{em} = 565$ nm) and 7-hydroxycoumarin-D-alanine (HADA; $\lambda_{em} = 450$ nm), were time-pulsed to record the chronological steps of cell wall growth in *Streptomyces venezuelae* (**Figure 5b**). This approach has also been used to monitor the production of peptidoglycans in single bacterial cells in different environments, (such as during antibiotic treatment[115] or transplantation[109], and to track dynamic interactions between different bacterial species. In a notable example of the latter, the predation cycle of *Bdellovibrio bacteriovorus* (a small bacterial species that preys on larger bacteria) and its prey *E. coli* was studied by super-resolution imaging of the multiplexed FDAAs HADA, BODIPY-3-amino-D-alanine (BADA; $\lambda_{em} = 512$ nm) and tetramethylrhodamine-3-amino-D-alanine (TADA; $\lambda_{em} = 565$ nm)[116].

Numerous studies have investigated the mechanism by which small and large FDAAs can be incorporated into peptidoglycans of multiple bacterial species. Most FDAAs behave as substrate analogues of solvent-accessible, periplasmic peptidoglycan transpeptidases, which are responsible for the assembly of the cell wall material during growth and division[117,118]. For example, HADA was used as an activity-based probe of peptidoglycan transpeptidases to monitor the spatial distribution of new cell wall versus old cell wall in live *Bacillus subtilis*[117]. Owing to their multiplexing capabilities, FDAAs can also be combined with fluorescently-tagged peptidoglycan biosynthetic proteins for multi-colour imaging studies. For example, metabolic incorporation of HADA was combined with the localization of key proteins in cell division (for example, FtsZ) in evolutionarily distinct rod-shaped bacterial species (such as *B. subtilis*[119] and *E. coli*[120]), and MreB was colocalized with an EDA-DA-labelled peptidoglycan ring during cell division in *Chlamydia trachomatis*[121]. Of note, these studies support a new paradigm in which bacterial division progresses directionally around the cell, in contrast to the previously described model of uniform cell division.

[H3] Tools for high-throughput screenings

However, macromolecular labelling of peptidoglycan structures with conventional FDAAs has constraints on temporal resolution due to the washing steps that are needed to reduce the background fluorescence from unincorporated free amino acids. These technical limitations prompted the design of a series of FDAAs that are non-fluorescent in low viscosity environments (that is, when the FDAA is moving freely in aqueous media) and fluoresce strongly when their intramolecular conformation is constrained (for example, after incorporation into peptidoglycans)[122] (**Figure 5c**). These so-called rotor-FDAAs (RfDAAs) consist of an electron-donating tetrahydroquinoline core structure coupled to electron-withdrawing carboxy-2-cyanovinyl groups to create efficient push–pull systems (**Figure 3c**). The resulting water-soluble amino acids Rf420DL (λ_{em} = 420 nm), Rf470DL (λ_{em} = 470 nm) and Rf490DL (λ_{em} = 490 nm) have enabled wash-free and real-time imaging of peptidoglycan biosynthesis, making possible the first high-throughput *in vitro* assay to probe the activity of peptidoglycan transpeptidases. The imaging capabilities of these FDAAs will facilitate their application in biological studies in which monitoring biological events in real time is of extreme importance, including morphogenesis and drug screening studies.

[H1] Genetically-encoded FlAAs

[H3] Encoding FlAAs into proteins

Proteins can be endowed with fluorescence properties by numerous methods, including by enzymatic and chemical site-specific labelling with optical reporters and by *in vitro* translation of proteins with fluorescent tags[123-127]. These methods have undoubtedly aided the study of complex biological processes but they do suffer from some limitations, including potential off-target fluorescence and restriction to *in vitro* assays. Bearing this in mind, the genetic encoding of FlAAs has rapidly evolved to enable the generation of artificial proteins that might overcome some of these challenges[128].

Site-specific incorporation of non-natural amino acids by genetic encoding in cells is generally achieved by assigning an amber stop codon to the non-natural amino acid and supplying the cells with an exogenous aminoacyl-tRNA synthetase (aaRS) and its corresponding transfer ribonucleic acid (tRNA). Importantly, the exogenous aaRS and tRNA must be orthogonal and not cross react with the endogenous aaRSs and tRNAs of the host cell. Then, the orthogonal aaRS specifically couples the orthogonal tRNA to the non-natural amino acid so that the ‘charged’ tRNA is used by the ribosome for site-specific incorporation of the non-natural amino acid (**Figure 6a**). Rapid advances in this technology have led to the efficient synthesis

of artificial proteins containing non-natural amino acids in live cells and animals[129-131]. To date, more than 150 non-natural amino acids, including bioorthogonal, photoreactive and photocaged amino acids, as well as amino acids with electron paramagnetic resonance (EPR), infrared (IR) and nuclear magnetic resonance (NMR) labels, have been genetically encoded in different organisms[132]. The development of organisms with synthetic genomes[133] and the evolution of quadruplet-codon-decoding ribosomes[134,135] will expand the applicability of new genetically-encoded FlAAs with different chemical structures and/or additional photophysical properties to label proteins.

[H3] Genetically-encoded FlAAs in prokaryotes

The FlAAs that have been genetically incorporated into proteins (**Figure 6b**) can be broadly classified by whether they are expressed in prokaryotic or in eukaryotic cells. Following the discovery and application of L-(7-hydroxycoumarin-4-yl)ethylglycine (CouA; $\lambda_{em.}$= 380 nm in neutral form; $\lambda_{em.}$= 450 nm in anionic form), acridon-2-ylalanine (Acd; $\lambda_{em.}$= 446 nm) and 4-biphenyl-l-phenylalanine (terphenylA; $\lambda_{em.}$= 342 nm) were developed as additional FlAAs for site-specific incorporation into proteins in *E. coli*[20,45,136]. In these cases, evolved mutant tyrosyl-aaRSs from the archaeon *Methanococcus jannaschii* (*Mj*TyrRS) were used to generate 'charged' tRNAs inside the cell, which resulted in successful ribosomal incorporation. Of note, the *Mj*TyrRS–*Mj*tRNA pair is orthogonal in prokaryotes and has been evolved for incorporation of several other non-natural amino acids.

CouA is synthesized by incorporating a β-keto ester in the amino acid side chain and then reacting it with resorcinol by Pechmann condensation[20]. The ease of synthesis and efficient incorporation of CouA into proteins has facilitated the widespread application of CouA in biological studies. As the emission spectrum of CouA overlaps with the excitation spectrum of cyan fluorescent protein (CFP), CouA was used to engineer a new CFP with a long Stokes shift of ~110 nm (ref.[137]). In this case, the precise incorporation of CouA at 20 Å from the CFP fluorophore results in efficient FRET so that CFP fluorescence (at 426 nm) is detected when CouA is excited (at 365 nm) (**Figure 6c**). Furthermore, the sensitivity of CouA to pH and polarity changes has also been exploited to produce artificial proteins with readouts that are sensitive to environmental conditions[138]. For example, this approach was used to investigate the difference in substrate specificity between the haloalkane dehalogenases DhaA and DbjA[139]. Steady-state and time-resolved fluorescence measurements of dehalogenase mutants containing CouA at specific sites revealed greater hydration in DbjA than in DhaA, which correlates with the substrate specificity of the two proteins. CouA has also been genetically

encoded into the transient receptor potential cation channel subfamily V member 1 (TRPV1) to study the dynamics of receptor activation, with changes in the fluorescence being correlated with the opening and closing of the channel on binding of capsaicin[140].

[H3] Genetically-encoded FIAAs in eukaryotes

Lysine derivatives of 7-hydroxycoumarin (Lys-Cou) have been genetically encoded into proteins using evolved mutants of *Methanosarcina barkeri* pyrrolysyl RS (PylRS)–$tRNA_{CUA}$[141]. The PylRS–tRNA pair[142,143] has become an extremely useful tool for genetic encoding of non-natural amino acids because it is orthogonal to prokaryotes and eukaryotes and therefore can be used in cells of evolutionarily diverse origins, including *E. coli*, *Saccharomyces cerevisiae*, mammalian cells and even whole intact organisms, such as the nematode worm *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster*[144]. ACD has also been encoded in fluorescent proteins for expression in prokaryotes. The high quantum yield (~95%, compared with 63% for CouA), efficient synthesis and high photostability of ACD are some of its key advantages.

Among the FlAAs that have been genetically encoded in eukaryotes, the most notable are dansyl and prodan derivatives, such as dansylalanine ($\lambda_{em.}$= 540 nm) and ANAP ($\lambda_{em.}$= 490 nm)[145-147] (**Figure 6b**). These FlAAs have been incorporated into proteins expressed in *S. cerevisiae* using evolved mutants of the *E. coli*-derived leucyl-RS–leucyl-$tRNA_{CUA}$ pair, which is orthogonal to those found in eukaryotic cells. A remarkable feature of these amino acids is that their emission maximum and quantum yield can vary drastically depending on the polarity of the medium; for example, the quantum yield of dansyl-based amino acids vary from 50% ($\lambda_{em.}$= 510 nm) in ethyl acetate to <10% ($\lambda_{em.}$= 578 nm) in water[148]; the emission maxima for ANAP are 490 nm in water and 420 nm in ethyl acetate. This sensitivity to the dielectric constant of the surrounding environment has been exploited to investigate the dynamics of different protein structures. For example, ANAP was used to map the regions of the voltage-gated potassium ion channel ion channel that are inaccessible to chemical labelling[149] (**Figure 6d**). First, the FlAA was genetically encoded into different positions in the ion channel sequence, either the extracellular and intracellular regions, and then electrophysiology and fluorescence measurements were performed concurrently to identify the regions of the protein that cooperatively lead to opening of the channel. Similarly, ANAP has been also used to examine the catalytic activity of a voltage-sensing phosphatase[150] and to gain insights into protein–protein interactions using FRET measurements, as in the case of the pro-apoptotic protein BAX and the anti-apoptotic protein HSP70 (ref.[151]). Different mutants of BAX with

ANAP encoded at three different positions in the protein were tested for binding to HSP70 fused to yellow fluorescent protein (YFP), so that FRET signals were detected only when the two proteins were in close proximity. This FRET signal was used as a reporter of the interaction between BAX and HSP70 and helped to discern the mechanism of small molecule-induced inhibition of this interaction. These examples highlight how genetically-encoded FlAAs can be used to gain insights into protein dynamics and folding, as well as the interactions between proteins and other biomolecules.

[H1] Conclusions and outlook

Advances in the chemical derivatization of heterocyclic building blocks and fluorophores have accelerated the design and synthesis of FlAAs with improved properties for chemical biology studies. To date, the unique reactivity of the indole moiety[152,153] has resulted in the synthesis of multiple tryptophan analogues by different transformations (such as Pd-catalysed couplings and C–H activation). Site-specific modification of other amino acids, including, amongst others, histidine[154,155], phenylalanine[49] and tyrosine[156,157], will help to enable other residues with complementary functionalities to be fluorescently labelled and, ultimately, lead to broader chemical diversity. Similarly, the emergence of new chemistry in the field of synthetic fluorophores will enable the construction of FlAAs with additional properties. These properties include not only physicochemical properties, such as small size[158,159] to reduce the effect of labelling in peptides and proteins, but also optical features for spectroscopy and imaging experiments, including near-infrared emission[160,161], super-resolution capabilities[162,163], photoactivatable behaviour[164,165] or suitability for multimodal imaging (for example, optoacoustics and positron emission tomography)[166,167]. A greater number and variety of FlAAs will create avenues to identify new roles for peptides and proteins in biological systems. FlAAs with longer emission wavelengths will improve the penetration depth that is required for *in vivo* imaging studies, and photoswitching FlAAs may allow researchers to image the localization and trafficking of proteins using super-resolution microscopy. New FlAAs might also find important applications in other areas of chemistry in which they may have been underexplored to date, such as the material sciences, where fluorescent building blocks embedded in supramolecular structures could provide dynamic readouts to characterize the formation and properties of new materials in real time, including nanofibres and peptide hydrogels[168,169].

For biological applications, the design of new methodologies in synthetic biology will facilitate the integration of new modular blocks into different macromolecular structures. An

area that will benefit substantially from these developments is genetically-encoded FlAAs. Although currently available genetically-encoded FIAAs hold huge potential for studying the dynamics, folding and biomolecular interactions of proteins, their number and types are limited. Most fluorophores absorb at very short wavelengths that require high energy excitation and lead to limited tissue penetration and poor signal-to-background ratios, [which precludes their use in whole organisms. Non-natural amino acids showing good spectral overlap with existing fluorophores or fluorescent proteins would be also extremely useful to build FRET pairs within proteins and thereby study dynamic conformational changes with increased resolution. The UV excitation wavelengths of current FlAAs also reduce their compatibility with photocaging groups, which are typically cleaved when illuminated at those wavelengths (for example, 365 nm) [170,171]. Therefore, synthetic routes that are able to red-shift the excitation and emission wavelengths[172] of non-natural FlAAs would have a remarkable impact on the design of new artificial proteins. However, this synthetic effort will need to be matched with advances in the biological counterpart, where the identification of new aaRS–tRNA pairs that can recognize and incorporate such molecules will be essential for the successful encoding of novel FlAAs into protein structures.

References

1 Chan, J., Dodani, S. C. & Chang, C. J. Reaction-based small-molecule fluorescent probes for chemoselective bioimaging. *Nat. Chem.* **4**, 973-984, (2012).

2 Lavis, L. D. & Raines, R. T. Bright building blocks for chemical biology. *ACS Chem. Biol.* **9**, 855-866 (2014).

3 Klymchenko, A. S. Solvatochromic and fluorogenic dyes as environment-sensitive probes: design and biological applications. *Acc. Chem. Res.* **50**, 366-375 (2017).

4 Park, S. J. et al. Mechanistic elements and critical factors of cellular reprogramming revealed by stepwise global gene expression analyses. *Stem Cell Res.* **12**, 730-741 (2014).

5 Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J. Cell. Comparative Physiol.* **59**, 223-239 (1962).

6 Tsien, R. Y. The green fluorescent protein. *Ann. Rev. Biochem.* **67**, 509-544 (1998).

7 Heim, R., Cubitt, A. B. & Tsien, R. Y. Improved green fluorescence. *Nature* **373**, 663-664 (1995).

8 Keppler, A. et al. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86-89 (2003).

9 Gautier, A. et al. An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15**, 128-136 (2008).

10 Los, G. V. et al. HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373-382 (2008).

11 Griffin, B. A., Adams, S. R. & Tsien, R. Y. Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269 (1998).

12 Martin, B. R., Giepmans, B. N. G., Adams, S. R. & Tsien, R. Y. Mammalian cell–based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat. Biotechnol.* **23**, 1308-1314 (2005).

13 Krueger, A. T. & Imperiali, B. Fluorescent amino acids: modular building blocks for the assembly of new tools for chemical biology. *ChemBioChem* **14**, 788-799 (2013).

14 Nitz, M., Mezo, A. R., Ali, M. H. & Imperiali, B. Enantioselective synthesis and application of the highly fluorescent and environment-sensitive amino acid 6-(2-dimethylaminonaphthoyl) alanine (DANA). *Chem. Commun.* 1912-1913 (2002).

15 Vázquez, M. E., Blanco, J. B. & Imperiali, B. Photophysics and biological applications of the environment-sensitive fluorophore 6-n,n-dimethylamino-2,3-naphthalimide. *J. Amer. Chem. Soc.* **127**, 1300-1306 (2005).

16 Socher, E. & Imperiali, B. FRET-Capture: a sensitive method for the detection of dynamic protein interactions. *ChemBioChem* **14**, 53-57 (2013).

17 Shults, M. D. & Imperiali, B. Versatile fluorescence probes of protein kinase activity. *J. Amer. Chem. Soc.* **125**, 14248-14249 (2003).

18 Venkatraman, P. et al. Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells. *Nat. Chem. Biol.* **3**, 222-228 (2007).

19 Vázquez, M. E., Rothman, D. M. & Imperiali, B. A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis. *Org. Biomol. Chem.* **2**, 1965-1966 (2004).

20 Wang, J., Xie, J. & Schultz, P. G. A genetically encoded fluorescent amino acid. *J. Amer. Chem. Soc.* **128**, 8738-8739 (2006).

21 Kielland, N., Vendrell, M., Lavilla, R. & Chang, Y.-T. Imaging histamine in live basophils and macrophages with a fluorescent mesoionic acid fluoride. *Chem. Commun.* **48**, 7401-7403 (2012).

22 Weiss, J. T. et al. Extracellular palladium-catalysed dealkylation of 5-fluoro-1-propargyl-uracil as a bioorthogonally activated prodrug approach. *Nat. Commun.* **5**, 3277-3285 (2014).
23 Ramil, C. P. & Lin, Q. Photoclick chemistry: a fluorogenic light-triggered in vivo ligation reaction. *Curr. Opin. Chem. Biol.* **21**, 89-95 (2014).
24 Devaraj, N. K. The future of bioorthogonal chemistry. *ACS Cent. Sci.* **4**, 952-959 (2018).
25 de Moliner, F., Kielland, N., Lavilla, R. & Vendrell, M. Modern synthetic avenues for the preparation of functional fluorophores. *Angew. Chem. Int. Ed.* **56**, 3758-3769 (2017).
26 Teale, F. W. & Weber, G. Ultraviolet fluorescence of the aromatic amino acids. *Biochem. J.* **65**, 476-482 (1957).
27 Ghisaidoobe, A. B. T. & Chung, S. J. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. *Int. J. Mol. Sci.* **15**, 22518-22538 (2014).
28 Merkel, L., Hoesl, M. G., Albrecht, M., Schmidt, A. & Budisa, N. Blue fluorescent amino acids as in vivo building blocks for proteins. *ChemBioChem* **11**, 305-314 (2010).
29 Talukder, P. et al. Cyanotryptophans as novel fluorescent probes for studying protein conformational changes and DNA–protein interaction. *Biochemistry* **54**, 7457-7469 (2015).
30 Hilaire, M. R. et al. Blue fluorescent amino acid for biological spectroscopy and microscopy. *Proc. Nat. Acad. Sci.* **114**, 6005-6009 (2017).
Small analogue of a naturally-occurring amino acid that behaves as a non-disruptive building block for the preparation of fluorescent peptides.
31 Zhang, K. et al. Synthesis and application of the blue fluorescent amino acid l-4-cyanotryptophan to assess peptide–membrane interactions. *Chem. Commun.* **55**, 5095-5098 (2019).
32 Winn, M., Francis, D. & Micklefield, J. De novo biosynthesis of “non-natural” thaxtomin phytotoxins. *Angew. Chem. Int. Ed.* **57**, 6830-6833 (2018).
33 Boville, C. E., Romney, D. K., Almhjell, P. J., Sieben, M. & Arnold, F. H. Improved synthesis of 4-cyanotryptophan and other tryptophan analogues in aqueous solvent using variants of TrpB from *Thermotoga maritima*. *J. Org. Chem.* **83**, 7447-7452 (2018).
34 Talukder, P., Chen, S., Arce, P. M. & Hecht, S. M. Efficient asymmetric synthesis of tryptophan analogues having useful photophysical properties. *Org. Lett.* **16**, 556-559 (2014).

35 Wen, J. et al. Highly N2-selective coupling of 1,2,3-triazoles with indole and pyrrole. *Chem. Eur. J.* **20**, 974-978 (2014).

36 Williams, T. J., Reay, A. J., Whitwood, A. C. & Fairlamb, I. J. S. A mild and selective Pd-mediated methodology for the synthesis of highly fluorescent 2-arylated tryptophans and tryptophan-containing peptides: a catalytic role for Pd^0 nanoparticles? *Chem. Commun.* **50**, 3052-3054 (2014).

37 Bartoccini, F., Bartolucci, S., Mari, M. & Piersanti, G. A simple, modular synthesis of C4-substituted tryptophan derivatives. *Org. Biomol. Chem.* **14**, 10095-10100 (2016).

38 Talukder, P. et al. Tryptophan-based fluorophores for studying protein conformational changes. *Bioorg. Med. Chem.* **22**, 5924-5934 (2014).

39 Chen, S. et al. Fluorescent biphenyl derivatives of phenylalanine suitable for protein modification. *Biochemistry* **52**, 8580-8589 (2013).

40 Chen, S. et al. Detection of dihydrofolate reductase conformational change by FRET using two fluorescent amino acids. *J. Amer. Chem. Soc.* **135**, 12924-12927 (2013).

41 Maity, J., Honcharenko, D. & Strömberg, R. Synthesis of fluorescent D-amino acids with 4-acetamidobiphenyl and 4-N,N-dimethylamino-1,8-naphthalimido containing side chains. *Tet. Lett.* **56**, 4780-4783 (2015).

42 Cheruku, P. et al. Tyrosine-derived stimuli responsive, fluorescent amino acids. *Chem. Sci.* **6**, 1150-1158 (2015).

A toolbox of tyrosine-based FlAAs with tunable emission and reversible pH and redox responses, showing potential for biosensing applications.

43 Bylińska, I., Guzow, K., Wójcik, J. & Wiczk, W. New non-proteinogenic fluorescent amino acids: benzoxazol-5-yl-alanine derivatives containing acetylene unit. Synthesis, spectral and photophysical properties. *J. Photochem. Photobiol. A: Chem.* **364**, 679-685 (2018).

44 Hoppmann, C., Alexiev, U., Irran, E. & Rück-Braun, K. Synthesis and fluorescence of xanthone amino acids. *Tet. Lett.* **54**, 4585-4587 (2013).

45 Speight, L. C. et al. Efficient synthesis and in vivo incorporation of acridon-2-ylalanine, a fluorescent amino acid for lifetime and förster resonance energy transfer/luminescence resonance energy transfer studies. *J. Amer. Chem. Soc.* **135**, 18806-18814 (2013).

46 Mendive-Tapia, L. et al. Preparation of a Trp-BODIPY fluorogenic amino acid to label peptides for enhanced live-cell fluorescence imaging. *Nat. Protoc.* **12**, 1588-1619 (2017).

47 Subiros-Funosas, R. et al. Fluorogenic Trp(redBODIPY) cyclopeptide targeting keratin 1 for imaging of aggressive carcinomas. *Chem. Sci.* **11**, 1368-1374 (2020).

48 Terrey, M. J., Holmes, A., Perry, C. C. & Cross, W. B. C–H Olefination of tryptophan residues in peptides: control of residue selectivity and peptide–amino acid cross-linking. *Org. Lett.* **21**, 7902-7907 (2019).

49 Wang, W., Lorion, M. M., Martinazzoli, O. & Ackermann, L. BODIPY peptide labeling by late-stage C(sp3)−H activation. *Angew. Chem. Int. Ed.* **57**, 10554-10558 (2018).

50 Schischko, A. et al. Late-stage peptide C–H alkylation for bioorthogonal C–H activation featuring solid phase peptide synthesis. *Nat. Commun.* **10**, 3553-3561 (2019).

51 Pereira, G., Vilaça, H. & Ferreira, P. M. T. Synthesis of new β-amidodehydroaminobutyric acid derivatives and of new tyrosine derivatives using copper catalyzed C–N and C–O coupling reactions. *Amino Acids* **44**, 335-344 (2013).

52 Bag, S. S., Jana, S. & Pradhan, M. K. Synthesis, photophysical properties of triazolyl-donor/acceptor chromophores decorated unnatural amino acids: Incorporation of a pair into Leu-enkephalin peptide and application of triazolylperylene amino acid in sensing BSA. *Bioorg. Med. Chem.* **24**, 3579-3595 (2016).

53 Benedetti, E., Veliz, A. B. E., Charpenay, M., Kocsis, L. S. & Brummond, K. M. Attachable solvatochromic fluorophores and bioconjugation studies. *Org. Lett.* **15**, 2578-2581 (2013).

54 Li, C. et al. Click chemistry to fluorescent amino esters: synthesis and spectroscopic studies. *Eur. J. Org. Chem.* **2010**, 2395-2405 (2010).

55 Ferreira, P. M. T., Monteiro, L. S., Pereira, G., Castanheira, E. M. S. & Frost, C. G. Synthesis of fluorescent alanines by a rhodium-catalysed conjugate addition of arylboronic acids to dehydroalanine derivatives. *Eur. J. Org. Chem.* **2013**, 550-556 (2013).

56 Hsu, Y.-P. et al. Full color palette of fluorescent D-amino acids for in situ labeling of bacterial cell walls. *Chem. Sci.* **8**, 6313-6321 (2017).

57 Gütschow, M. & Häußler, D. Fluorescently labeled amino acids as building blocks for bioactive molecules. *Synthesis* **48**, 245-255 (2015).

58 Kuru, E., Tekkam, S., Hall, E., Brun, Y. V. & Van Nieuwenhze, M. S. Synthesis of fluorescent D-amino acids and their use for probing peptidoglycan synthesis and bacterial growth in situ. *Nat. Protoc.* **10**, 33-52 (2014).

59 Katritzky, A. R., Ozcan, S. & Todadze, E. Synthesis and fluorescence of the new environment-sensitive fluorophore 6-chloro-2,3-naphthalimide derivative. *Org. Biomol. Chem.* **8**, 1296-1300 (2010).

60 Esteves, C. I. C., Silva, A. M. F., Raposo, M. M. M. & Costa, S. P. G. Unnatural benz-X-azolyl asparagine derivatives as novel fluorescent amino acids: synthesis and photophysical characterization. *Tetrahedron* **65**, 9373-9377 (2009).

61 Yokoo, H., Kagechika, H, Ohsaki, A. & Hirano, T. A polarity-sensitive fluorescent amino acid and its incorporation into peptides for the ratiometric detection of biomolecular interactions. *ChemPlusChem* **84**, 1716-1719 (2019).

62 Shukla, L., Moodie, L. W. K., Kindahl, T. & Hedberg, C. Synthesis and spectroscopic properties of fluorinated coumarin lysine derivatives. *J. Org. Chem.* **83**, 4792-4799 (2018).

63 Bag, S. S. & De, S. Isothiocyanyl alanine as a synthetic intermediate for the synthesis of thioureayl alanines and subsequent aminotetrazolyl alanines. *J. Org. Chem.* **82**, 12276-12285 (2017).

64 Mohite, A. R. & Bhat, R. G. Enantiopure synthesis of side chain-modified α-amino acids and 5-cis-alkylprolines. *J. Org. Chem.* **77**, 5423-5428 (2012).

65 Fowler, L. S., Ellis, D. & Sutherland, A. Synthesis of fluorescent enone derived α-amino acids. *Org. Biomol. Chem.* **7**, 4309-4316 (2009).

66 Navo, C. D. et al. Cell-penetrating peptides containing fluorescent d-cysteines. *Chem. Eur. J.* **24**, 7991-8000 (2018).

67 Wörner, S., Rönicke, F., Ulrich, A. S. & Wagenknecht, H-A. 4-Aminophthalimide acids as small and environment-sensitive fluorescent probes for transmembrane peptides. *ChemBioChem* **21**, 618-622 (2020).

68 Xiang, Z. & Wang, L. Enantiospecific synthesis of genetically encodable fluorescent unnatural amino acid l-3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid. *J. Org. Chem.* **76**, 6367-6371 (2011).

69 Zhao, Y., Pirrung, M. C. & Liao, J. A fluorescent amino acid probe to monitor efficiency of peptide conjugation to glass surfaces for high density microarrays. *Mol. BioSyst.* **8**, 879-887 (2012).

70 Arribat, M., Rémond, E., Clément, S., Lee, A. V. D. & Cavelier, F. Phospholyl(borane) amino acids and peptides: stereoselective synthesis and fluorescent properties with large stokes shift. *J. Am. Chem. Soc.* **140**, 1028-1034 (2018).

71 Strizhak, A. V. et al. Two-color fluorescent L-amino acid mimic of tryptophan for probing peptide–nucleic acid complexes. *Bioconjugate Chem.* **23**, 2434-2443 (2012).

First report of a tyrosine-based FlAA exhibiting excited-state intramolecular proton transfer and hydration sensitive dual emission.72 Postupalenko, V. Y. et al. Dual-fluorescence L-amino acid reports insertion and orientation of melittin peptide in cell membranes. *Bioconjugate Chem.* **24**, 1998-2007 (2013).

73 Sholokh, M. et al. Fluorescent amino acid undergoing excited state intramolecular proton transfer for site-specific probing and imaging of peptide interactions. *J. Phys. Chem. B* **119**, 2585-2595 (2015).

74 Taki, M., Yamazaki, Y., Suzuki, Y. & Sisido, M. Introduction of a highly photodurable and common-laser excitable fluorescent amino acid into a peptide as a FRET acceptor for protease cleavage detection. *Chem. Lett.* **39**, 818-819 (2010).

75 Bell, J. D. et al. Synthesis and photophysical properties of benzotriazole-derived unnatural α-amino acids. *J. Org. Chem.* **84**, 10436-10448 (2019).

76 Gilfillan, L., Artschwager, R., Harkiss, A. H., Liskamp, R. M. J. & Sutherland, A. Synthesis of pyrazole containing α-amino acids via a highly regioselective condensation/aza-Michael reaction of β-aryl α,β-unsaturated ketones. *Org. Biomol. Chem.* **13**, 4514-4523 (2015).

77 Harkiss, A. H., Bell, J. D., Knuhtsen, A., Jamieson, A. G. & Sutherland, A. Synthesis and fluorescent properties of β-pyridyl α-amino acids. *J. Org. Chem.* **84**, 2879-2890 (2019).

78 Bell, J. D. et al. Conformationally rigid pyrazoloquinazoline α-amino acids: one- and two-photon induced fluorescence. *Chem. Commun.* **56**, 1887-1890 (2020).

79 Häußler, D. & Gütschow, M. Synthesis of a fluorescent-labeled bisbenzamidine containing the central (6,7-dimethoxy-4-coumaryl)alanine building block. *Heteroatom Chem.* **26**, 367-373 (2015).

80 Koopmans, T., van Haren, M., van Ufford, L. Q., Beekman, J. M. & Martin, N. I. A concise preparation of the fluorescent amino acid l-(7-hydroxycoumarin-4-yl) ethylglycine and extension of its utility in solid phase peptide synthesis. *Bioorg. Med. Chem.* **21**, 553-559 (2013).

81 Moodie, L. W. K., Chammaa, S., Kindahl, T. & Hedberg, C. Palladium-mediated approach to coumarin-functionalized amino acids. *Org. Lett.* **19**, 2797-2800 (2017).

82 Fernandez, A., Thompson, E. J., Pollard, J. W., Kitamura, T. & Vendrell, M. A fluorescent activatable AND-gate chemokine CCL2 enables in vivo detection of metastasis-associated macrophages. *Angew. Chem. Int. Ed.* **58**, 16894-16898 (2019).

83 Joshi, P. N. & Rai, V. Single-site labeling of histidine in proteins, on-demand reversibility, and traceless metal-free protein purification. *Chem. Commun.* **55**, 1100-1103 (2019).

84 Cheng, M. H. Y., Savoie, H., Bryden, F. & Boyle, R. W. A convenient method for multicolour labelling of proteins with BODIPY fluorophores via tyrosine residues. *Photochem. Photobiol. Sci.* **16**, 1260-1267 (2017).

85 Zhao, C. et al. Searching for the optimal fluorophore to label antimicrobial peptides. *ACS Comb. Sci.* **18**, 689-696 (2016).

86 Vendrell, M. et al. Biotin ergopeptide probes for dopamine receptors. *J. Med. Chem.* **54**, 1080-1090 (2011).

87 Palomo, J. M. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. *RSC Adv.* **4**, 32658-32672 (2014).

88 Sainlos, M., Iskenderian, W. S. & Imperiali, B. A general screening strategy for peptide-based fluorogenic ligands: probes for dynamic studies of PDZ domain-mediated interactions. *J. Am. Chem. Soc.* **131**, 6680-6682 (2009).

Pioneering work on solvatochromic phthalimide amino acids and their integration into peptide structures to study dynamic protein-protein interactions.

89 Loving, G. & Imperiali, B. A versatile amino acid analogue of the solvatochromic fluorophore 4-n,n-dimethylamino-1,8-naphthalimide: a powerful tool for the study of dynamic protein interactions. *J. Am. Chem. Soc.* **130**, 13630-13638 (2008).

90 Wang, W. et al. A fluorogenic peptide probe developed by in vitro selection using tRNA carrying a fluorogenic amino acid. *Chem. Commun.* **50**, 2962-2964 (2014).

91 Wang, W. et al. Fluorogenic enhancement of an in vitro-selected peptide ligand by replacement of a fluorescent group. *Anal. Chem.* **88**, 7991-7997 (2016).

92 Heru, C., Jurgen, S., Tino, Z. & Horst, A. Fluorescent analogues of the insect neuropeptide helicokinin I: synthesis, photophysical characterization and biological activity. *Prot. Pept. Lett.* **17**, 431-436 (2010).

93 Manandhar, Y. et al. In vitro selection of a peptide aptamer that changes fluorescence in response to verotoxin. *Biotechnol. Lett.* **37**, 619-625 (2015).

94 Newton, L. D., Pascu, S. I., Tyrrell, R. M. & Eggleston, I. M. Development of a peptide-based fluorescent probe for biological heme monitoring. *Org. Biomol. Chem.* **17**, 467-471 (2019).

95 Zhao, C., Mendive-Tapia, L. & Vendrell, M. Fluorescent peptides for imaging of fungal cells. *Arch. Biochem. Biophys.* **661**, 187-195 (2019).

96 Mendive-Tapia, L. et al. Spacer-free BODIPY fluorogens in antimicrobial peptides for direct imaging of fungal infection in human tissue. *Nat. Commun.* **7**, 10940-10948 (2016).

First report of Trp-BODIPY as a fluorogenic amino acid to non-invasively label peptides for live-cell and *ex vivo* tissue imaging

97 Akram, A. R. et al. Enhanced avidity from a multivalent fluorescent antimicrobial peptide enables pathogen detection in a human lung model. *Sci. Rep.* **9**, 8422-8431 (2019).

98 Subiros-Funosas, R. et al. A Trp-BODIPY cyclic peptide for fluorescence labelling of apoptotic bodies. *Chem. Commun.* **53**, 945-948 (2017).

99 Ge, J., Li, L. & Yao, S. Q. A self-immobilizing and fluorogenic unnatural amino acid that mimics phosphotyrosine. *Chem. Commun.* **47**, 10939-10941 (2011).

100 K.C, T. B. et al. Wash-free and selective imaging of epithelial cell adhesion molecule (EpCAM) expressing cells with fluorogenic peptide ligands. *Biochem. Biophys. Res. Commun.* **500**, 283-287 (2018).

101 Maeno, T. et al. Targeted delivery of fluorogenic peptide aptamers into live microalgae by femtosecond laser photoporation at single-cell resolution. *Sci. Rep.* **8**, 8271-8279 (2018).

102 Egan, A. J. F., Cleverley, R. M., Peters, K., Lewis, R. J. & Vollmer, W. Regulation of bacterial cell wall growth. *FEBS J.* **284**, 851-867 (2017).

103 Dörr, T., Moynihan, P. J. & Mayer, C. Bacterial cell wall structure and dynamics. *Front. Microbiol.* **10**, 2051 (2019).

104 Radkov, A. D., Hsu, Y.-P., Booher, G. & VanNieuwenhze, M. S. Imaging bacterial cell wall biosynthesis. *Ann. Rev. Biochem.* **87**, 991-1014 (2018).

105 Hsu, Y.-P., Booher, G., Egan, A., Vollmer, W. & VanNieuwenhze, M. S. D-Amino acid derivatives as in situ probes for visualizing bacterial peptidoglycan biosynthesis. *Acc. Chem. Res.* **52**, 2713-2722 (2019).

106 de Pedro, M. A., Quintela, J. C., Höltje, J. V. & Schwarz, H. Murein segregation in *Escherichia coli*. *J. Bacteriol.* **179**, 2823-2834 (1997).

107 Cava, F., de Pedro, M. A., Lam, H., Davis, B. M. & Waldor, M. K. Distinct pathways for modification of the bacterial cell wall by non-canonical D-amino acids. *EMBO J.* **30**, 3442-3453 (2011).

108 Lupoli, T. J. et al. Transpeptidase-mediated incorporation of D-amino acids into bacterial peptidoglycan. *J. Am. Chem. Soc.* **133**, 10748-10751 (2011).

109 Siegrist, M. S. et al. D-amino acid chemical reporters reveal peptidoglycan dynamics of an intracellular pathogen. *ACS Chem. Biol.* **8**, 500-505 (2013).

110 Kuru, E. et al. In situ probing of newly synthesized peptidoglycan in live bacteria with fluorescent D-amino acids. *Angew. Chem. Int. Ed.* **51**, 12519-12523 (2012).

Development of a toolbox of fluorescent D-amino acids for labelling peptidoglycans and monitor bacterial cell wall growth.

111 Liechti, G. W. et al. A new metabolic cell-wall labelling method reveals peptidoglycan in *Chlamydia trachomatis*. *Nature* **506**, 507-510 (2013).

112 Hudak, J. E., Alvarez, D., Skelly, A., von Andrian, U. H. & Kasper, D. L. Illuminating vital surface molecules of symbionts in health and disease. *Nat. Microbiol.* **2**, 17099-17106 (2017).

113 Wang, W. et al. Assessing the viability of transplanted gut microbiota by sequential tagging with D-amino acid-based metabolic probes. *Nat. Commun.* **10**, 1317-1323 (2019).

114 Hu, F. et al. Visualization and in situ ablation of intracellular bacterial pathogens through metabolic labeling. *Angew. Chem. Int. Ed.* doi:10.1002/anie.201910187 (2019).

115 Kuru, E. Mechanisms of incorporation for D-amino acid probes that target peptidoglycan biosynthesis. *ACS Chem. Biol.* **14**, 2745-2756 (2019).

116 Kuru, E. et al. Fluorescent D-amino-acids reveal bi-cellular cell wall modifications important for *Bdellovibrio bacteriovorus* predation. *Nat. Microbiol.* **2**, 1648-1657 (2017).

117 Morales Angeles, D. et al. Pentapeptide-rich peptidoglycan at the *Bacillus subtilis* cell-division site. *Mol. Microbiol.* **104**, 319-333 (2017).

118 Baranowski, C. et al. Maturing *Mycobacterium smegmatis* peptidoglycan requires non-canonical crosslinks to maintain shape. *eLife* **7**, e37516 (2018).

119 Bisson-Filho, A. W. et al. Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *Science* **355**, 739-743 (2017).

120 Yang, X. et al. GTPase activity–coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. *Science* **355**, 744-747 (2017).

121 Liechti, G. et al. Pathogenic *Chlamydia* lack a classical sacculus but synthesize a narrow, mid-cell peptidoglycan ring, regulated by MreB, for cell division. *PLOS Pathogens* **12**, e1005590 (2016).

122 Hsu, Y.-P. et al. Fluorogenic D-amino acids enable real-time monitoring of peptidoglycan biosynthesis and high-throughput transpeptidation assays. *Nat. Chem.* **11**, 335-341 (2019).

Rotor fluorogenic D-amino acids for real-time visualization of transpeptidase reactions and high-throughput screening of antibacterial drugs.

123 Rodriguez, E. A. et al. The growing and glowing toolbox of fluorescent and photoactive proteins. *Trends Biochem. Sci.* **42**, 111-129 (2017).

124 Jing, C. & Cornish, V. W. Chemical tags for labeling proteins inside living cells. *Acc. Chem. Res.* **44**, 784-792 (2011).

125 Spicer, C. D. & Davis, B. G. Selective chemical protein modification. *Nat. Commun.* **5**, 4740-4753 (2014).

126 Kajihara, D. et al. FRET analysis of protein conformational change through position-specific incorporation of fluorescent amino acids. *Nat. Methods* **3**, 923-929 (2006).

127 Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **6**, 779-790 (2011).

128 Lang, K. & Chin, J. W. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **114**, 4764-4806 (2014).

129 Young, D. D. & Schultz, P. G. Playing with the molecules of life. *ACS Chem. Biol.* **13**, 854-870 (2018).

130 Chin, J. W. Expanding and reprogramming the genetic code of cells and animals. *Ann. Rev. Biochem.* **83**, 379-408 (2014).

131 Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53-60 (2017).

132 Dumas, A., Lercher, L., Spicer, C. D. & Davis, B. G. Designing logical codon reassignment – expanding the chemistry in biology. *Chem. Sci.* **6**, 50-69 (2015).

133 Fredens, J. et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **569**, 514-518 (2019).

134 Wang, K. et al. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* **6**, 393-403 (2014).

135 Sachdeva, A., Wang, K., Elliott, T. & Chin, J. W. Concerted, rapid, quantitative, and site-specific dual labeling of proteins. *J. Am. Chem. Soc.* **136**, 7785-7788 (2014).

136 Lampkowski, J. S., Uthappa, D. M. & Young, D. D. Site-specific incorporation of a fluorescent terphenyl unnatural amino acid. *Bioorg. Med. Chem. Lett.* **25**, 5277-5280 (2015).

137 Kuhn, S. M., Rubini, M., Müller, M. A. & Skerra, A. Biosynthesis of a fluorescent protein with extreme pseudo-stokes shift by introducing a genetically encoded non-natural amino acid outside the fluorophore. *J. Am. Chem. Soc.* **133**, 3708-3711 (2011).

138 Lacey, V. K. et al. A fluorescent reporter of the phosphorylation status of the substrate protein STAT3. *Angew. Chem. Int. Ed.* **50**, 8692-8696 (2011).

139 Amaro, M. et al. Site-specific analysis of protein hydration based on unnatural amino acid fluorescence. *J. Am. Chem. Soc.* **137**, 4988-4992 (2015).

140 Steinberg, X. et al. Conformational dynamics in TRPV1 channels reported by an encoded coumarin amino acid. *eLife* **6**, e28626 (2017).

141 Luo, J. et al. Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. *J. Am. Chem. Soc.* **136**, 15551-15558 (2014).

142 Wan, W., Tharp, J. M. & Liu, W. R. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta – Prot. Proteom.* **1844**, 1059-1070 (2014).

143 Srinivasan, G., James, C. M. & Krzycki, J. A. Pyrrolysine encoded by UAG in archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**, 1459-1462 (2002).

144 Elliott, T. S. et al. Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat. Biotechnol.* **32**, 465-472 (2014).

145 Summerer, D. et al. A genetically encoded fluorescent amino acid. *Proc. Nat. Acad. Sci.* **103**, 9785-9789 (2006).

146 Lee, H. S., Guo, J., Lemke, E. A., Dimla, R. D. & Schultz, P. G. Genetic incorporation of a small, environmentally sensitive, fluorescent probe into proteins in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **131**, 12921-12923 (2009).

Seminal work on the genetic encoding of the environmentally-sensitive amino acid ANAP into protein structures.

147 Chatterjee, A., Guo, J., Lee, H. S. & Schultz, P. G. A genetically encoded fluorescent probe in mammalian cells. *J. Am. Chem. Soc.* **135**, 12540-12543 (2013).

148 Chen, R. F. Fluorescence of dansyl amino acids in organic solvents and protein solutions. *Arch. Biochem. Biophys.* **120**, 609-620 (1967).

149 Kalstrup, T. & Blunck, R. Dynamics of internal pore opening in Kv channels probed by a fluorescent unnatural amino acid. *Proc. Nat. Acad. Sci.* **110**, 8272-8277 (2013).

The incorporation of ANAP into proteins enabled the investigation of the gating mechanism of voltage-gated potassium channels.

150 Sakata, S., Jinno, Y., Kawanabe, A. & Okamura, Y. Voltage-dependent motion of the catalytic region of voltage-sensing phosphatase monitored by a fluorescent amino acid. *Proc. Nat. Acad. Sci.* **113**, 7521-7526 (2016).

151 Park, S.-H., Ko, W., Lee, H. S. & Shin, I. Analysis of protein–protein interaction in a single live cell by using a FRET system based on genetic code expansion technology. *J. Am. Chem. Soc.* **141**, 4273-4281 (2019).

152 Preciado, S. et al. Synthesis and biological evaluation of a post-synthetically modified Trp-based diketopiperazine. *MedChemComm* **4**, 1171-1174 (2013).

153 Mendive-Tapia, L. et al. New peptide architectures through C–H activation stapling between tryptophan–phenylalanine/tyrosine residues. *Nat. Commun.* **6**, 7160-7168 (2015).

154 Chen, X. et al. Histidine-specific peptide modification via visible-light-promoted C–H alkylation. *J. Am. Chem. Soc.* **141**, 18230-18237 (2019).

155 Peciak, K., Laurine, E., Tommasi, R., Choi, J.-W. & Brocchini, S. Site-selective protein conjugation at histidine. *Chem. Sci.* **10**, 427-439 (2019).

156 Ban, H., Gavrilyuk, J. & Barbas, C. F. Tyrosine bioconjugation through aqueous ene-type reactions: a click-like reaction for tyrosine. *J. Am. Chem. Soc.* **132**, 1523-1525 (2010).

157 Tilley, S. D. & Francis, M. B. Tyrosine-selective protein alkylation using π-allylpalladium complexes. *J. Am. Chem. Soc.* **128**, 1080-1081 (2006).

158 Benson, S. et al. SCOTfluors: small, conjugatable, orthogonal, and tunable fluorophores for in vivo imaging of cell metabolism. *Angew. Chem. Int. Ed.* **58**, 6911-6915 (2019).

159 Su, L. et al. Cu(I)-catalyzed 6-endo-dig cyclization of terminal alkynes, 2-bromoaryl ketones, and amides toward 1-naphthylamines: applications and photophysical properties. *J. Am. Chem. Soc.* **141**, 2535-2544 (2019).

160 Mellanby, R. J. et al. Tricarbocyanine *N*-triazoles: the scaffold-of-choice for long-term near-infrared imaging of immune cells in vivo. *Chem. Sci.* **9**, 7261-7270 (2018).

161 Cosco, E. D. et al. Flavylium polymethine fluorophores for near- and shortwave infrared imaging. *Angew. Chem. Int. Ed.* **56**, 13126-13129 (2017).
162 Tang, J. et al. Single-atom fluorescence switch: a general approach toward visible-light-activated dyes for biological imaging. *J. Am. Chem. Soc.* **141**, 14699-14706 (2019).
163 Zheng, Q. et al. Rational design of fluorogenic and spontaneously blinking labels for super-resolution imaging. *ACS Cent. Sci.* **5**, 1602-1613 (2019).
164 Raymo, F. M. Photoactivatable synthetic fluorophores. *Phys. Chem. Chem. Phys.* **15**, 14840-14850 (2013).
165 Zhang, Y., Tang, S., Sansalone, L., Baker, J. D. & Raymo, F. M. A photoswitchable fluorophore for the real-time monitoring of dynamic events in living organisms. *Chem. Eur. J.* **22**, 15027-15034 (2016).
166 Hendricks, J. A. et al. Synthesis of [^{18}F]BODIPY: bifunctional reporter for hybrid optical/positron emission tomography imaging. *Angew. Chem. Int. Ed.* **51**, 4603-4606 (2012).
167 Zhou, E. Y., Knox, H. J., Liu, C., Zhao, W. & Chan, J. A conformationally restricted aza-BODIPY platform for stimulus-responsive probes with enhanced photoacoustic properties. *J. Am. Chem. Soc.* **141**, 17601-17609 (2019).
168 Onogi, S. et al. In situ real-time imaging of self-sorted supramolecular nanofibres. *Nat. Chem.* **8**, 743-752 (2016).
169 Beesley, J. L. & Woolfson, D. N. The de novo design of α-helical peptides for supramolecular self-assembly. *Curr. Opin. Biotechnol.* **58**, 175-182 (2019).
170 Davis, L. & Greiss, S. Genetic encoding of unnatural amino acids in *C. elegans*. *Meth. Mol. Biol.* **1728**, 389-408 (2018).
171 Bridge, T. et al. Site-specific encoding of photoactivity in antibodies enables light-mediated antibody–antigen binding on live cells. *Angew. Chem. Int. Ed.* **58**, 17986-17993. (2019).
172 Ren, T.-B. et al. A general method to increase stokes shift by introducing alternating vibronic structures. *J. Am. Chem. Soc.* **140**, 7716-7722 (2018).

Acknowledgements

E. K. acknowledges funding of a Fellowship from the Life Sciences Research Foundation. A.S. acknowledges funding from the Wellcome Trust (204593/Z/16/Z) and the Biotechnology and

Biological Research Sciences Council (BB/R004692/1). M.V. acknowledges funding from an ERC Consolidator Grant (771443). The authors thank S. Shaikh for useful comments and technical support with the graphical illustrations.

Author contributions

All authors contributed equally to the preparation of this manuscript.

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Chemistry thanks A. Klymchenko and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note

Figure legends

Figure 1. **Applications of FlAAs in chemical biology.** The design and synthesis of new fluorescent amino acids (FlAAs) has created multiple opportunities for *in vitro* biomolecular recognition assays (part **a**), solid-phase peptide synthesis (SPPS) of fluorescent peptides for optical imaging (part **b**), genetic encoding of fluorescent building blocks into proteins (part **c**) and microbiology experiments to study bacterial growth (part **d**). Images in part **b** reproduced with permission from ref. 96, Springer Nature Limited. Image in part **d** reproduced with permission from ref. 112, Springer Nature Limited.

Figure 2. Representative non-natural FlAAs developed in the past decade. Chemical structures and basic spectral properties of fluorescent analogues of tryptophan (part **a**) and fluorescent amino acids (FlAAs) obtained by appending fluorophores to amino acids using a broad range of chemical reactions (part **b**). ϕ_F, fluorescence quantum yield; Pd, palladium; Rh, rhodium; CuAAC, copper-catalysed azide–alkyne cycloaddition.

Figure 3. Synthetic schemes for the *de novo* preparation of non-natural FlAAs. **a |** Design of flavone-based fluorescent amino acids (FlAAs) building on tyrosine derivatization[71-73]. **b |** Pyrazoloquinazolines derived from L-aspartic acid as one-photon and two-photon fluorophores[78]. **c |** Synthesis of rotor fluorogenic D-amino acids (RfDAAs) for labelling peptidoglycan structures in bacterial cells[122]. Boc, *tert*-butoxycarbonyl; DCM, dichloromethane; DDQ, 2,3-dichloro-5,6-dicyano-1,4-benzoquinone; DIPEA, *N*,*N*-diisopropylethylamine; DMF, *N*,*N*-dimethyl formamide; Fmoc-OSu, *N*-(9-fluorenylmethyloxycarbonyl)succinimide; TFA, trifluoroacetic acid ; THF, tetrahydrofuran; Trt-Cl, triphenylmethyl chloride; TSTU, *N*,*N*,*N*′,*N*′-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate.

Figure 4. Bioactive fluorescent peptides for structural studies and optical imaging. **a |** Schematic representation of the interaction between the HIV-1 nucleocapsid peptide W37 (labelled with the flavone-based fluorescent amino acid (FlAA) M3HFaa) and oligonucleotides. The graph shows the fluorescence emission spectra of the W37–M3HFaa peptide on binding to different nucleic acids. **b |** Schematic illustration of the binding between a fluorescent peptide containing the FlAA 4-DMAP and a PDZ domain. The graph shows the fluorescence emission spectra of the 4-DMAP-containing peptide in the unbound state and bound to the PDZ domain. **c |** Illustration of the pH-(low) insertion peptide (pHLIP) containing

4-cyanotryptophan (which acts as a FRET donor, and DiO-stained large unilamellar vesicles (LUVs), which act as FRET acceptors, for the study of peptide–membrane interactions. The graph depicts the FRET measurements taken at 512 nm with increasing concentrations of labelled pHLIP. **d |** Chemical structure of the Trp-BODIPY-labelled cyclic peptide PAF26 and live-cell confocal imaging of *Aspergillus fumigatus* over time. Scale bar: 2.5 μm. Multi-photon fluorescence microscopy images of *ex vivo* human lung tissue after incubation with RFP-expressing *A. fumigatus* (red) and the Trp-BODIPY-labelled PAF26 peptide (green). Second harmonic generation images are shown in cyan. Scale bar: 10 μm. Fluorescence lifetime images depicting autofluorescent tissue structures (white arrows) and a PAF26-stained *A. fumigatus* hypha (yellow arrows). Scale bar: 20 μm. **e |** pER peptide containing a self-immobilizing coumarin FlAA that mimics phosphotyrosine (pTyr) to report endogenous activity of protein tyrosine phosphatases (PTPs). The graph depicts time-dependent fluorescence measurements of the pER peptide before and after ultraviolet (UV) irradiation (365 nm) in HeLa cell lysates. Fluorescence microscopy images of HeLa cells after incubation with pER and UV irradiation to visualize intracellular PTP activity (green: pER peptide, red: commercial tracker). Part **a** reproduced with permission from ref. 73, American Chemical Society. Part **b** was reproduced with permission from ref. 88, American Chemical Society. Part **c** reproduced with permission from ref. 31, The Royal Society of Chemistry. Part **d** reproduced with permission from ref. 96, Springer Nature Limited. Part **e** reproduced with permission from ref. 99, The Royal Society of Chemistry.

Figure 5. **Fluorescent D-amino acids for studying bacterial growth. a |** Peptidoglycans are polymers of L- and D-amino acids and sugars, and a major component of bacterial cell walls. The peptidoglycan layer is considerably thicker in Gram-positive than in Gram-negative bacteria. **b |** Chemical structures and optical properties of 'clickable' and fluorescent D-amino acids (FDAAs) used for labelling peptidoglycan structures in Gram-positive and Gram-negative bacteria. **c |** Time-lapse multicolour microscopy images of the bacterium *Streptomyces venezuelae* by pulsed-incubation with NBD-3-amino-D-alanine (NADA; green), tetramethylrhodamine-D-lysine (TDL; red), and 7-hydroxycoumarin-D-alanine (HADA; blue). **d |** Schematic representation of the fluorescence activation of the rotor FDAA Rf470DL on peptidoglycan labelling and fluorescence microscopy images of *Bacillus subtilis* depicting peptidoglycan labelling using Rf470DL (red) or HADA (cyan) before and after washing. Part **c** reproduced with permission from ref. 110, Wiley VCH. Part **d** reproduced with permission from ref. 122, Springer Nature Limited.

Figure 6. Genetically encoded site-specific incorporation of FlAAs in live cells. a | Non-natural amino acids are site-specifically incorporated into proteins by assigning an amber stop codon (UAG) to non-natural amino acids and supplying the cells with an exogenous aminoacyl-tRNA synthetase (aaRS) and its corresponding tRNA, which charges the tRNA with the non-natural amino acid. **b |** Chemical structures of genetically-encoded fluorescent amino acids (FlAAs) and their optical properties. **c |** Adjusting the photophysical properties of a fluorescent protein. Site-specific incorporation of L-(7-hydroxycoumarin-4-yl)ethylglycine (CouA) at position 39 in cyan fluorescent protein (CFP), resulting in efficient FRET between CouA and CFP. When CouA (in blue) is excited at 360 nm, emission from the CFP fluorophore (in cyan) at 476 nm is observed. **d |** Mapping ion channel dynamics. 3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (ANAP) is site-specifically incorporated into both extracellular and intracellular positions (black circles) in a voltage-gated potassium ion channel. Voltage-dependent changes in the fluorescence emission of ANAP revealed information about the regions of the ion channel that move cooperatively on opening of the channel. Red spheres in the structure are potassium ions. Part **c** adapted with permission from ref. 137, American Chemical Society. Part **d** adapted with permission from ref. 149, National Academy of Sciences USA.

ToC blurb

Fluorescent amino acids are widely used as building blocks for non-perturbative labelling of peptides and proteins. This Review covers recent advances in the design and synthesis of FlAAs with bespoke optical properties for different applications in biological studies.